9-28-2017

The Dark Side of Interpolated Testing: Frequent Switching Between Retrieval and Encoding Impairs New Learning

Sara D. Davis
Iowa State University, sddavis@iastate.edu

Jason C.K. Chan
Iowa State University, ckchan@iastate.edu

Miko M. Wilford
University of Massachusetts Lowell

The Dark Side of Interpolated Testing: Frequent Switching Between Retrieval and Encoding Impairs New Learning

Abstract
Practicing retrieval can improve the updating or modification of existing knowledge. When students need to update their existing knowledge, performing retrieval practice on the first set of materials often strengthens learning of the next set. However, Davis and Chan (2015)reported that interpolated testing can sometimes impair new learning. Here, we examined whether frequently switching between retrieval of previously learned material and encoding of new material can disrupt learning of the new material. In the current experiment, participants either switched between restudying originally learned items and new learning or between retrieving originally learned items and new learning, and we varied the frequency with which task switching occurred. We found that interpolating retrieval, but not restudy, with new learning impaired new learning. These results are consistent with the idea that retrieval practice and encoding rely on different cognitive processes, and intermixing them can exert a cost.

Keywords
testing effect, interpolated testing, test-potentiated learning, retrieval, task-switching

Disciplines
Psychology

Comments
This is an accepted manuscript of an article published as Davis, Sara D., Jason CK Chan, and Miko M. Wilford. "The Dark Side of Interpolated Testing: Frequent Switching Between Retrieval and Encoding Impairs New Learning." *Journal of Applied Research in Memory and Cognition* 6, no. 4 (2017): 434-441. doi: 10.1016/j.jarmac.2017.07.002. Posted with permission.

The Dark Side of Interpolated Testing:

Frequent Switching Between Retrieval and Encoding Impairs New Learning

Sara D. Davis

Jason C.K. Chan

Iowa State University

Miko Wilford

University of Massasschusett Lowell

Abstract

Practicing retrieval can improve the updating or modification of existing knowledge. When students need to update their existing knowledge, performing retrieval practice on the first set of materials often strengthens learning of the next set. However, Davis and Chan (2015) reported that interpolated testing can sometimes impair new learning. Here, we examined whether frequently switching between retrieval of previously learned material and encoding of new material can disrupt learning of the new material. In the current experiment, participants either switched between restudying originally learned items and new learning or between retrieving originally learned items and new learning, and we varied the frequency with which task switching occurred. We found that interpolating retrieval, but not restudy, with new learning impaired new learning. These results are consistent with the idea that retrieval practice and encoding rely on different cognitive processes, and intermixing them can exert a cost.

Keywords: testing effect, interpolated testing, test-potentiated learning, retrieval, task-switching

The Dark Side of Interpolated Testing:

Frequent Switching Between Retrieval and Encoding Impairs New Learning

Effective learning involves not only retaining information (e.g., students may see a picture of the hippocampus and must remember its shape or location in association with its name), but updating existing knowledge with new information (e.g., the functions of various areas of the hippocampus). Knowledge updating is particularly important in learning STEM concepts at the university level. For example, students may need to first master the simpler concepts in general biology and later update their knowledge by learning the more complex processes like gene expression (Jensen, Kummer, & Banjoko, 2013). As such, an important goal in educational research is to identify techniques that can aid in both retention and knowledge updating.

Retrieval practice (or testing) is one of the most effective techniques for boosting learning. In the present paper, we focus on the finding that testing can potentiate new learning that occurs later (Pastotter & Bauml, 2014l; Szpunar, McDermott, & Roediger, 2008; Wissman & Rawson, 2011). This benefit of testing has important educational implications, as performing retrieval practice on previously learned information may facilitate knowledge updating. To this end, researchers have argued that interspersing a lecture with brief memory tests can facilitate learning in the classroom (Szpunar, Khan, & Schacter, 2013). However, and of critical interest to the present study, interspersing retrieval practice trials into an encoding phase has also been shown to impair new learning (Davis & Chan, 2015; Finn, 2017; Finn & Roediger, 2013).

In a series of experiments, Finn and Roediger (2013) demonstrated that performing retrieval practice can impair new learning in a memory updating paradigm.

Specifically, participants first studied a set of face-name pairs – the original learning (*OL*) items. During the intermediate phase, participants either restudied a face-name pair or recalled the name when given the face (before receiving feedback). Following this restudy or test trial, participants must update their knowledge of the face by learning the profession for that face – the new learning (*NL*) items. Surprisingly, initial testing of the OL association impaired subsequent learning of the NL association.

In a subsequent study, Davis and Chan (2015) argued that testing impaired new learning in Finn and Roediger's (2013) paradigm because intermixing retrieval practice and new learning trials might have encouraged participants to prioritize restudying the OL items ahead of learning the NL items. This bias could occur because the test reveals to participants the difficulty associated with learning the OL item, and because the corrective feedback for the OL item is presented just before the NL item was presented. Consequently, participants might "borrow time" from the NL trial to study the OL item. Critically, this bias in relearning the OL items is absent when participants restudy the OL association, because they would not realize the difficulty associated with learning the OL association through restudy trials. Moreover, this bias in relearning the OL item is also absent in procedures that have typically demonstrated test-potentiated new learning, in which the retrieval practice trials and the new learning trials are presented in separate trial blocks (Pastötter & Bäuml, 2014; Weinstein, McDermott, & Szpunar, 2011; Szpunar, Khan, & Schacter, 2013; Wissman & Rawson, 2011).

These opposite effects of retrieval on new learning pose interesting questions for educational practice. For example, how often and when should instructors interpolate questions in a lecture to maximize their benefits and minimize their costs? From a theoretical perspective, it is important to examine the processes by which interpolated retrieval impairs new learning. Although Davis and Chan (2015) provided evidence that supports the metacognitive bias account, their manipulations, which were aimed at removing the bias on relearning the OL items, never resulted in the test-potentiated new learning effect. Instead, these manipulations either reduced or eliminated the test-impaired new learning effect (see their Experiments 2, 3, and 5).

In the present study, we consider the possibility that repeatedly switching between retrieval (relative to restudy) of OL items and encoding of NL items can impair new learning (see also Finn, 2017). This hypothesis stems from the cognitive control literature, wherein participants who alternate between two incompatible task sets often show impaired performance on one or both tasks (Allport, Styles, & Hsieh, 1994; Pashler, Johnston, & Ruthruff, 2001). Switching between two tasks requires participants to exert top-down control to reconfigure the task set (Baddeley et al., 2001; Logan & Gordon, 2001; Meiran, 1996), and the effort required to switch between incompatible tasks incurs a *switch cost* (Monsell, 2003; Rogers & Monsell, 1995; Sudevan & Taylor, 1987). For example, participants might be shown a series of digits, and must switch between indicating whether the number is even or odd in one trial and whether it is less than or greater than five in the next trial. Repeatedly switch between these incompatible tasks often degrades performance. The cost of task switching is manifested by poorer accuracy and/or slower reaction times on *switch trials* (e.g., trials preceded by a different task) relative to *stay trials* (e.g., consecutive trials of the same task). Applying this logic to the present context, requiring learners to switch between retrieval practice of OL items and encoding of NL items might elicit a switch cost, which could in turn impair new learning.

This proposition is supported by data suggesting that encoding and retrieval are subserved by different neural mechanisms. Tulving (1983) proposed the concept of *retrieval mode*, an active process of episodic remembering that occurs when individuals think back to previous experiences, which involves different processing demands than encoding. This idea is buttressed by neuroimaging evidence that shows that retrieval is right-lateralized and encoding is left-lateralized in the prefrontal cortex (e.g., Duzel, 2000; Tulving et al., 1994), and that retrieval and encoding employ different networks of the hippocampal formation (Duncan, Tompary, & Davachi, 2014).

A closer examination of Davis and Chan's (2015) data provides preliminary support for the task switch hypothesis. For example, Davis and Chan found that increasing the feedback duration associated with the OL item for a few seconds before encoding of the NL item decreased the magnitude of the test-impaired new learning effect (Experiment 2). This is similar to an established finding in the task-switching literature, in which switch costs are reduced when preparation time is increased (Meiran, 1996). Moreover, presenting the OL items and NL items in separate trial blocks led to test-potentiated new learning (Experiment 4). It is important to note here that the metacognitive bias account proposed by Davis and Chan and the task-switching account proposed here are not mutually exclusive, and both processes could contribute to the test-impaired new learning effect.

To test the possibility that switching between retrieval of studied information and encoding of new information may impair new learning, we parametrically manipulated task-switch frequency. Task-switch frequency refers to the number of times participants alternated between retrieval practice/restudy trials and new learning. We hypothesized that the negative impact of task switching on new learning would be exacerbated with more frequent switches between retrieval practice and new learning (i.e., encoding) operations, but not between restudy and new learning operations. It is important to note, however, that the way that students learn materials like those used in the current study may not generalize to every type of material that a student may encounter in the classroom. The present paradigm uses materials that can be represented as AB-AD associations, and task switching may affect new learning differently when students learn other types of associations (e.g., AB-CD).

Method

Participants

Four hundred undergraduate students at Iowa State University participated in this experiment for course credit. Twenty-two participants were eliminated because English was not their primary language and two were eliminated because they failed to follow instructions, yielding data from 376 participants for analysis. The number of participants in each condition is reported in Table 1. We determined minimum sample sizes a priori based on the difference in final NL performance between the restudy and test conditions in Davis and Chan's (2015) Experiments 1 and 4. In order to calculate the necessary number of subjects, we averaged these effect sizes (which were .57 and .81 respectively) to yield an average expected effect size of .69 In G*Power software, using an alpha value of .05, a power value of .80 and the expected effect size of .69, the analysis indicated that 34 participants per condition would be necessary to detect this same effect size in the current experiment. Data collection continued through the end of the semester, which resulted in several conditions exceeding this minimum value (see Table 1).

Design

The study used a 2 (Item Type: OL items or NL items) x 2 (Interpolated Task: Restudy or Test) x 5 (Task-Switch Frequency: (0, 3, 7, 19, or 39 total switches) mixed design. Item type was manipulated within-subjects, and interpolated task and task-switch frequency were manipulated between-subjects.

Materials

Study materials were the twenty face-name-profession triads from Davis and Chan (2015). Ten faces were female (M_{age} = 35.5 years) and 10 were male (M_{age} = 35.7 years). All faces were Caucasian with a neutral, closed-mouth expression (Minear & Park, 2004). The names (OL items) and professions (NL items) were taken from Finn and Roediger's (2013) stimuli. Each face was randomly paired with a gender-appropriate name, and professions were randomly assigned to each face-name pair.

Procedure

A graphical representation of the procedure is presented in Figure 1.[1] The initial encoding and final test phases were the same for all conditions. During the encoding phase, the 20 faces were presented for 5 s each and the name was presented below each face. A 500 ms inter-stimulus interval separated each face-name pair. After the study phase, participants completed 10 math problems for 6 s each. In the subsequent interpolated phase, participants either restudied the face-name association (i.e., an OL item) before they learned the face-profession association (i.e., an NL item) or they

[1] In each phase (study, interpolated restudy/testing, new learning, and final tests), the faces were presented in a new random order.

practiced retrieval for the face-name association before learning the face-profession association. For each restudy trial, the face and the correct name were presented together for 5 s. For each retrieval practice trial, the face was presented with a prompt for the participant to recall the name. The retrieval practice trials were self-paced[2], and participants pressed the "Enter" key to advance to the feedback presentation, which showed the correct name below the face for 2 s. For all participants, the new learning trials containing the face with the profession then appeared for 5 s. This procedure ensured that every participant had the same amount of time to learn the NL items (i.e., the face-profession association).

During the interpolated phase, all participants restudied or practiced retrieval on the 20 OL items and learned 20 NL items. Consequently, there were a total of 40 trials in the interpolated phase. Across the five switch frequency conditions, participants switched between restudy/retrieval practice and new learning on every trial, every two trials, every five trials, every ten trials, or every 20 trials. In the last condition, restudy/retrieval practice and new learning were separated by an instruction screen, which yielded zero switches. While switching between instruction encoding and learning the NL items could technically be construed as task switching, Rogers and Monsell (1995) found that allowing preparation time between blocks of tasks poses no switch cost. Thus, we do not identify the first encoding trial in this condition as a switch trial.

---

[2] On average, participants took about 6.4 s to perform each retrieval practice trial. This duration was significantly longer than the 5 s allotted to participants in the restudy condition, $t(192) = 8.322$, $p < .001$, $d = 0.599$. However, as will be clear later, interpolated task did not have a main effect on final recall performance for the OL items. But most importantly, regardless of how long participants spent on the retrieval or restudy of the OL items, they had the same amount of time (5 s) to study the NL items.

The other conditions contained 39, 19, 7, and 3 switches, respectively, which occurred when participants learned an NL item following restudy or retrieval practice of an OL item (See Figure 1).

After the interpolated task and the new learning phase, all participants completed math problems for 60 s. They then took a final test for all 20 professions, which assessed new learning (NL), and then they were tested on all 20 names, which assessed original learning (OL). For each self-paced test trial, the face appeared along with a prompt to recall the appropriate item.

Results

All data and experiment materials are available at the Open Science Framework Repository, and can be found at the following web address: https://osf.io/c429j/

The Impact of Switch Frequency on Final Test Performance

We conducted separate 2 (Interpolated Task: Restudy or Test) x 5 (Task-Switch Frequency: 0, 3, 7, 19, or 39 total switches) between-subjects ANOVAs for the dependent variables of original learning (OL) and new learning (NL). The mean recall proportions that contribute to these ANOVAs are presented in Figure 2 with the corresponding 95% confidence intervals. We first consider the results for original learning. As shown in the top panel of Figure 2, neither the main effects of interpolated task, switch frequency, nor their interaction significantly affected final recall of the OL items, all Fs < 1.81, ps > .126, η_p^2s < .02. That is, task-switch frequency did not influence the effects of retrieval practice (relative to restudy) on final recall performance of the OL items. We now consider the results for new learning. Here, the ANOVA revealed a marginally significant main effect for task-switch frequency, $F(4, 366) = 2.03$, $p = .089$, $\eta_p^2 = .022$, and a nonsignificant main effect for interpolated task, $F(1, 366) =$

2.32, $p = .128$, $\eta_p^2 = .006$. Most importantly, as shown in the bottom panel of Figure 2, there was a crossover interaction between interpolated task type and task-switch frequency, $F(4, 366) = 9.58$, $p < .001$, $\eta_p^2 = .10$. Specifically, increasing task-switch frequency reduced NL recall for participants in the retrieval practice condition (i.e., a switch frequency cost), $F(4, 187) = 10.81$, $p < .001$, $\eta_p^2 = .19$, but it nonsignificantly increased NL recall for participants in the restudy condition, $F(4, 179) = 1.97$, $p = .101$, $\eta_p^2 = .04$.

To further examine the influence of task switching on new learning, we compared performance for the NL items between the restudy and retrieval practice conditions at each level of task-switch frequency. In the 0-switch condition, recalling the OL item had a large benefit on new learning relative to restudying the OL item, $t(76) = 3.67$, $p < .001$, $d = .83$, but this benefit was substantially smaller in the 3-switch condition, $t(69) = 1.75$, $p = .085$, $d = .41$. In the 7-switch condition, OL retrieval moderately *impaired* new learning relative to OL restudy, $t(75) = -2.08$, $p = .041$, $d = -.48$, and this impairment effect increased in the 19-switch condition, $t(78) = -2.67$, $p = .009$, $d = -.60$, and was substantially larger in the 39-switch condition, $t(74) = -4.49$, $p < .001$, $d = -1.03$. Thus, increasing task-switch frequency first reduced the benefit of retrieval practice on new learning, and then reversed and magnified the impairment.

Final Recall for NL Items Encoded on Switch and Stay Trials

We also examined final recall of NL items that were encoded on a switch or a stay trial. Data from the 0-switch and 39-switch conditions were omitted from this analysis as the former condition contained only stay trials and the latter condition contained only switch trials.

Switch trials were new learning trials that followed a restudy or retrieval practice trial, whereas stay trials were new learning trials that followed another new learning trial. To illustrate, in the 19-switch condition, participants alternated between two retrieval practice/restudy trials and two new learning trials in succession. Thus, half of the new learning trials occurred immediately after a restudy or retrieval practice trial, and these were classified as switch trials, with the remaining half classified as stay trials. In the 3- and 7-switch conditions, there were fewer switch trials than stay trials. Specifically, in the 7-switch condition for instance, only four NL items were encoded on a switch trial and the remaining 16 NL items were encoded on a stay trial. In the following analyses, we present results from two separate 3 (Task-Switch Frequency: 3, 7, 19) x 2 (Trial Type: Switch vs. Stay) ANOVAs. The first examined results for participants in the restudy condition, and the second examined results for participants in the retrieval practice condition. These data and their corresponding 95% confidence intervals are presented in Figure 3.

For participants who restudied the OL items, there was a main effect of task switch frequency, $F(2, 108) = 4.29$, $p = .016$, $\eta_p^2 = .07$. This main effect indicates that more frequent switching between restudy trials and new learning trials improved new learning.[3] However, there was no evidence of a switch cost, with NL items encoded on switch (M= .51) and stay trials (M = .54) exhibiting similar performance during the final

[3] Although ample research has demonstrated the negative influence of intermixing different tasks on cognitive performance, a related literature has shown that interleaving test trials that rely on the same type of processing can produce an interleaving benefit (Kornell & Bjork, 2008; Rohrer, Dedrick, & Burgess, 2014; Tauber, Dunlosky, Rawson, Wahlheim, & Jacoby, 2013).

test, $F(1, 108) = 1.54$, $p = .217$, $\eta_p^2 = .01$. Lastly, the interaction between task-switch frequency and trial type was not significant, $F < 1$. Thus, this analysis provides evidence that switching between tasks that require similar cognitive processes does not incur a cost and, indeed, might benefit new learning.

Unlike the participants in the restudy condition, participants in the retrieval practice condition exhibited poorer final recall performance for *NL items that were encoded on switch trials* ($M = .44$) than those encoded on stay trials ($M = .52$), $F(1, 114) = 8.16$, $p = .005$, $\eta_p^2 = .07$. Moreover, neither the main effect of task-switch frequency nor the interaction were significant, *F*s < 1.36. Taken together, the negative influences of switch frequency (i.e., deficient new learning when more switches occurred) and switch trials (i.e., deficient new learning on switch trials relative to stay trials) provide converging evidence for the idea that encoding and retrieval rely on different cognitive task sets, and alternating between them impaired new learning in the present experiment.

Discussion

In the present experiment, we showed that having learners repeatedly switch between retrieval and encoding can reverse the typically positive influence of retrieval practice on new learning. We demonstrated this dark side of interpolated testing in two important ways. First, as task switching became more frequent, the beneficial effects of retrieval practice on new learning decreased and then became detrimental. Critically, task-switch frequency had different effects on new learning depending on whether participants practiced retrieval or restudied the OL items before encoding the NL items. Only when participants switched between *retrieval* (and not restudy) and new learning did more frequent switching impair new learning. Second, NL items encoded on switch trials were recalled less often on the final test than those encoded on stay trials, regardless of task-switch frequency. Together, these results provide converging evidence for our hypothesis that switching between retrieval of previously studied information and encoding of new information can impair new learning.

These results have important implications for the conceptualization of retrieval and encoding processes. Tulving (1983) described the cognitive state involved in episodic retrieval as “retrieval mode,” which entails different cognitive and neurological demands from encoding (Duncan et al., 2014; Evans et al., 2015; Tulving et al., 1994). In the present context, we believe that testing impaired new learning because successful encoding of an NL item on switch trials required reconfiguration of the task set from retrieval of an OL item. This task set reconfiguration between retrieval (rather than restudying) and new encoding could have interfered with the encoding processes in the new learning trial. This task set reconfiguration for new learning was unnecessary when participants had just restudied an OL item on the previous trial.

A potential criticism of the method employed here is that by manipulating the frequency of task switching, we also manipulated the spacing between the presentation of each face-name (OL) association and the corresponding face-profession (NL) association. There is a great deal of research has shown that spaced presentations are beneficial to learning relative to massed presentations (see Cepeda et al., 2006, and Delaney, Verkoeijen, & Spirgel, 2010, for reviews). However, differential spacing cannot account for the present findings for one important reason. The spacing literature has found that both spaced study (e.g., Cuddy & Jacoby, 1982; Greene, 1989) *and* spaced retrieval (see Cepeda et al., 2006 for a review) enhance learning. Given that only the retrieval practice conditions, and not the restudy conditions, “benefitted” from spaced presentations (e.g.

new learning was enhanced when switching was less frequent, or the lag between face presentations was large), it seems unlikely that the spacing effect on its own could predict the interactions observed in this experiment.

Davis and Chan (2015) previously argued that a metacognitive bias toward relearning the OL items (in lieu of learning the NL items) could explain test-impaired new learning. Data from the present experiment, however, support the task-switching explanation. It is important to note again that the two accounts are not mutually exclusive, and that the present experiment was not designed to exclude the metacognitive bias account. In fact, intermixing retrieval practice with new learning trials might impair new learning *both* because it biases participants towards re-learning the OL items and because switching enacts a cost on NL encoding.

From an educational perspective, the current experiment demonstrates that retrieval practice does not universally benefit knowledge updating, and may in some cases impair students' ability to learn new information. One might argue that the materials used in the present study represent only a subset of the kinds of materials that students learn, and that memory updating of this nature will not always be necessary to learn advanced concepts. That is certainly true, but these materials do provide a reasonable analog for the type of AB-AD learning that frequently occurs in the classroom. In many educational contexts, knowledge must be scaffolded onto previously-learned concepts, which is similar to how participants may learn our materials. To return to our previous example, students might have to associate the name of the hippocampus with its picture (similar to associating a name with a face), as well as its function (similar to associating the same face with a profession).

Instructors might also be unlikely to alternate between administering retrieval practice in the classroom and lecturing new material to the same degree that was implemented in the present study. However, switching between retrieval and encoding had a negative impact on new learning (i.e., poorer NL recall for items encoded on switch vs. stay trials) at *every frequency* tested in our experiment, which suggests that a cost on new learning may occur even when retrieval practice is implemented sparingly during a lecture. Further, task switching incurred an immediate cost to encoding of new learning items introduced right after retrieval, which suggests that even one retrieval practice episode (e.g., a clicker quiz question) could negatively impact learning of new material that immediately follows. A practical solution to this problem is to minimize the number of switches that occur in any given learning episode, for example, by administering retrieval practice at either the beginning or the end of the lectures (Weinstein, Nunes, & Karpicke, 2016). Alternatively, instructors can spend time reviewing the relevant test material (i.e., restudy) prior to teaching new material to ease the transition of task sets (Rogers & Monsell, 1995). The viability of this approach in the classroom is an important question that awaits future research.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General, 130*(4), 641-657. doi: 10.1037/0096-3445.130.4.641

Buckner, R. L. (1996). Beyond HERA: Contributions of specific prefrontal brain areas to long-term memory retrieval. *Psychonomic Bulletin & Review, 3*(2), 149-158.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354-380.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning & Verbal Behavior, 21*(4), 451-467.

Davis, S.D., & Chan, J.C.K. (2015). Studying on borrowed time: How does testing impair new learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* doi: 10.1037/xlm0000126

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation*, *53*, 63–147.

Duncan, K., Tompary, A., & Davachi, L. (2014). Associative encoding and retrieval are predicted by functional connectivity in distinct hippocampal area CA1 pathways.

The Journal of Neuroscience, 34(34), 11188-11198. doi: 10.1523/JNEUROSCI.0521-14.2014

Düzel, E. (2000). When, where, what: The electromagnetic contribution to the WWW of brain activity during recognition. *Acta Psychologica, 105*(2-3), 195-210.

Finn, B. (2017). A framework of episodic updating: An account of memory updating after retrieval. *Psychology of Learning and Motivation*, *67*, 173–211.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(3), 371-377.

Finn, B., & Roediger, H.L. (2013). Interfering effects of retrieval in learning new information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(6), 1665-1681. doi:10.1037/a0032377

Jensen, J., Kummer, T., & Banjoko, A. (2013). Assessing the Effects of Prior Conceptions on Learning Gene Expression. *Journal of College Science Teaching, 42*(4), 82-91.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin, 136*(5), 849-874. doi: 10.1037/a0019842.

Kornell, N., & Bjork, R.A. (2008). Learning concepts and categories: Is spacing the "enemy of induction?". *Psychological Science, 19*(6), 585-592. doi: 10.1111/j.1467-9280.2008.02127.x

Logan, G.D., & Gordon, R.D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*(2), 393-434. doi: 10.1037/0033-295X.108.2.393

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1423-1442. doi:10.1037/0278-7393.22.6.1423

Minear, M. & Park, D.C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers, 36,* 630-633. doi:10.3758/BF03206543

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134-140. doi: 10.1016/S1364-6613(03)00028-7

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52*, 629-651.

Pastötter, B., & Bäuml, K.H.T. (2014). Retrieval practice enhances new learning: The forward effect of testing. *Frontiers in Psychology, 5*.

Rogers, R.D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231. doi: 10.1037/0096-3445.124.2.207

Rohrer, D., Dedrick, R.F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review, 21*(5), 1323-1330. doi:10.3758/s13423-014-0588-3

Sudevan, P., & Taylor, D.A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 89-103. doi: 10.1037/0096-1523.13.1.89

Szpunar, K.K., Khan, N.Y., & Schacter, D.L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences of the United States of America, 110*(16), 6313-6317.

Tauber, S.K., Dunlosky, J., Rawson, K.A., Wahlheim, C.N., & Jacoby, L.L. (2013). Self-regulated learning of a natural category: Do people interleave or block exemplars during study? *Psychonomic Bulletin & Review, 20*(2), 356-363.

Tulving E. (1983). *Elements of Episodic Memory*. Oxford: Clarendon.

Tulving, E., Kapur, S., Craik, F.I.M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences, 91,* 2016-2020.

Weinstein, Y., McDermott, K. B., & Szpunar, K. K. (2011). Testing protects against proactive interference in face–name learning. *Psychonomic Bulletin & Review, 18*(3), 518-523. doi: 10.3758/s13423-011-0085-x

Weinstein, Y., Nunes, L. D., & Karpicke, J. D. (2016). On the placement of practice questions during study. *Journal of Experimental Psychology: Applied, 22*(1), 72-84.

Wissman, K.T., Rawson, K.A., & Pyc, M.A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin &*

Review, 18(6), 1140-1147. doi:http://dx.doi.org.proxy.lib.iastate.edu/10.3758/s13423-011-0140-7

Table 1

Number of Participants in Each Between-Subjects Condition

	Task-Switch Frequency				
	0	***3***	***7***	***19***	***39***
Restudy	38	34	40	37	37
Test with Feedback	40	37	37	43	35

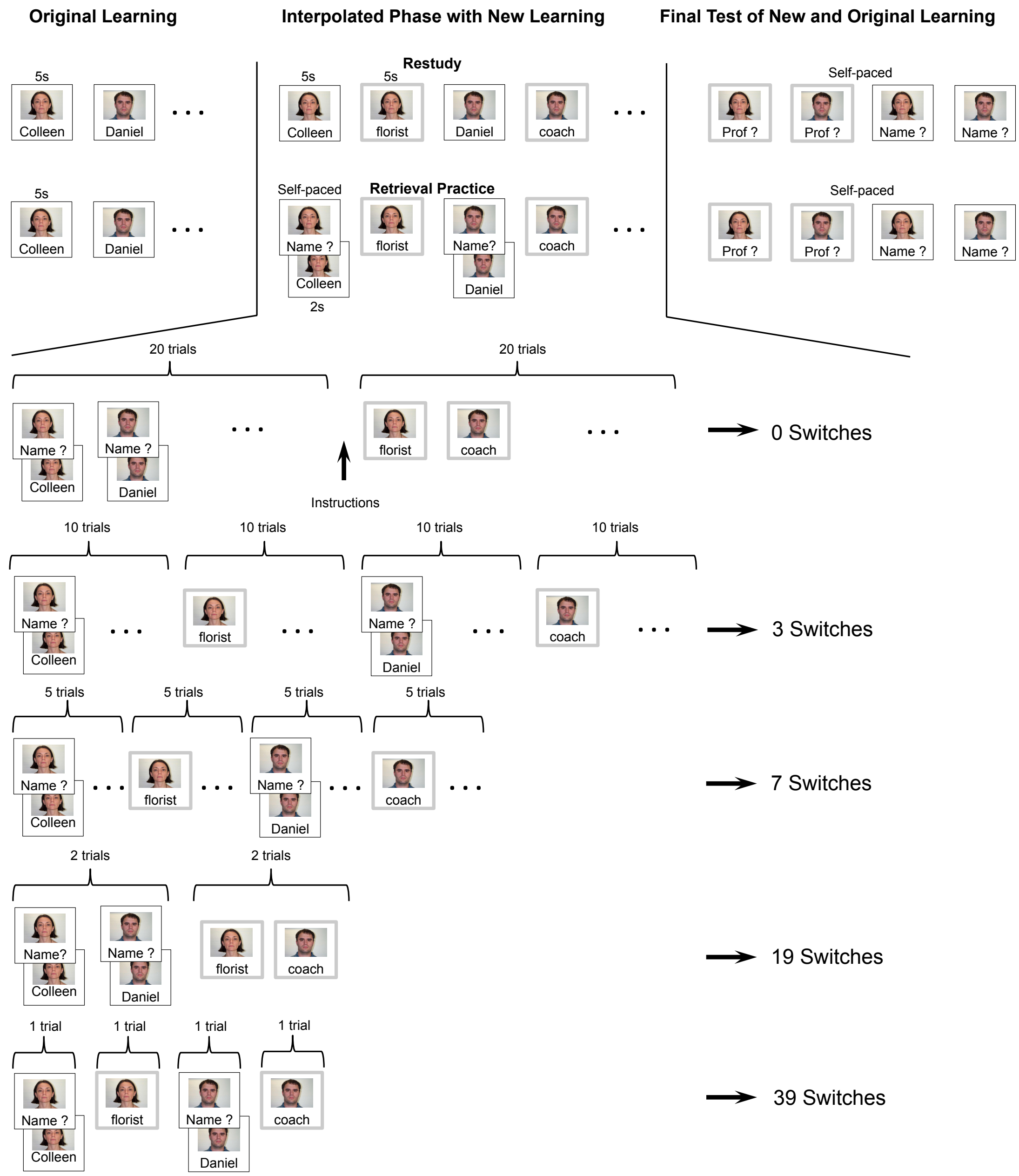

Figure 1. A graphical representation of the experimental procedure. New learning trials are shown in gray boxes. Note that the presentation order of faces with their associations was randomized in each phase.

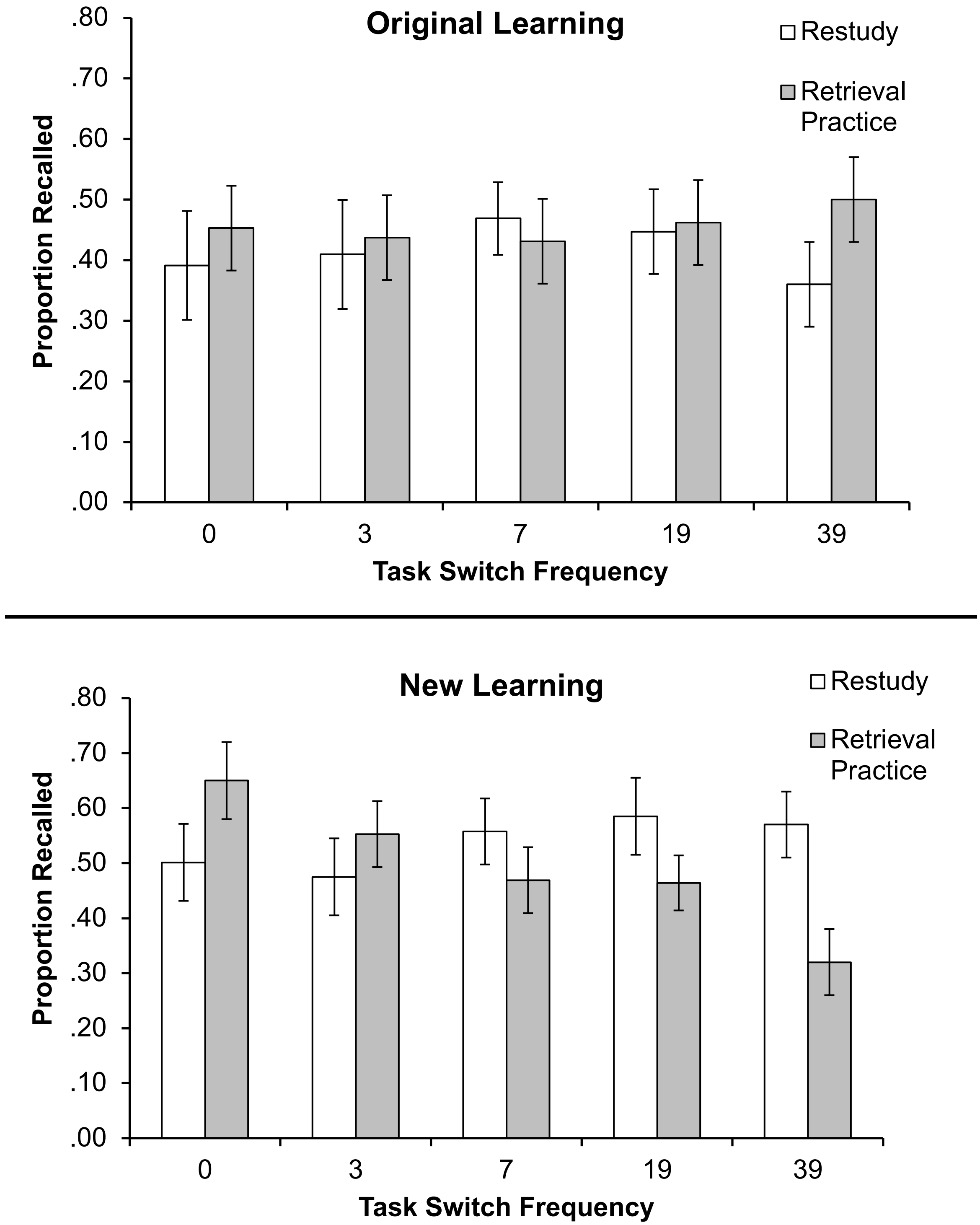

Figure 2. Final test performance for original learning (top panel) and new learning (bottom panel) as a function of task switch frequency for the restudy and retrieval practice conditions. Error bars are descriptive 95% confidence intervals.

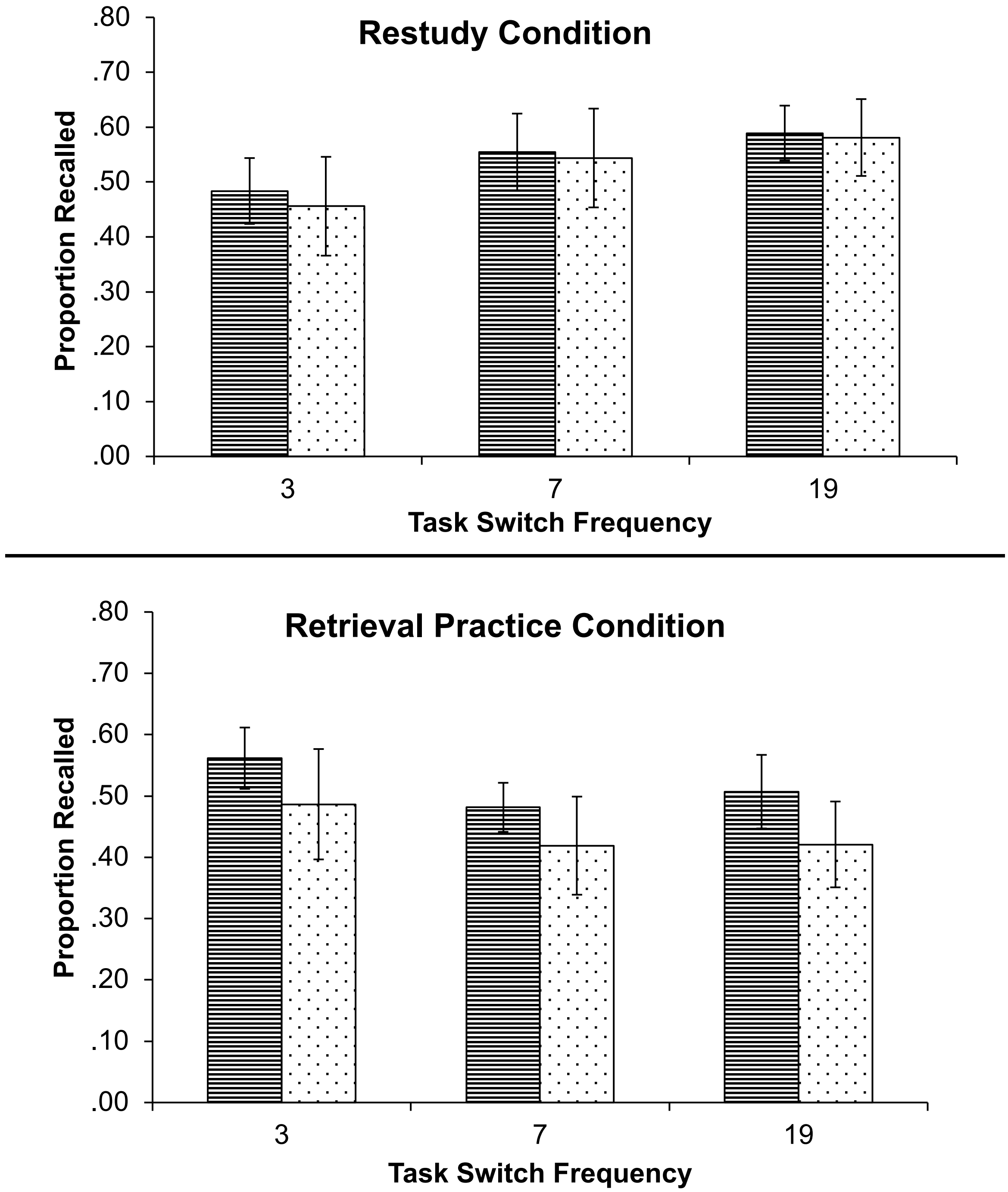

Figure 3. Final test performance for NL items on switch and stay trials based on task switch frequency for the restudy (top panel) and test with feedback (bottom panel) conditions (excluding the 0- and 39-switch conditions). Error bars are descriptive 95% confidence intervals.